第46卷 第1期
2023年1月

干旱区地理
ARID LAND GEOGRAPHY

Vol. 46 No. 1
Jan. 2023

基于机器学习和多光谱遥感的银川平原土壤盐分预测

魏慧敏[1]， 贾科利[1]， 张 旭[1]， 张俊华[2]

（1. 宁夏大学地理科学与规划学院，宁夏 银川 750021；2. 宁夏大学生态环境学院西北土地退化与生态恢复国家重点实验室培育基地，宁夏 银川 750021）

摘 要：快速获取区域土壤盐渍化程度信息，对于盐渍化治理与生态环境保护具有重要意义。以银川平原为研究区，以盐分影响因子和盐分指数分别作为输入参数，建立支持向量机（SVM），BP神经网络（BPNN）和贝叶斯神经网络（BNN）3种土壤盐分预测模型，选取最佳模型进行研究区不同深度的土壤盐渍化预测。结果表明：（1）0~20 cm土壤盐分预测模型中基于影响因子变量组的BNN模型效果最佳，决定系数（R^2）为0.618，均方根误差（RMSE）为2.986；20~40 cm土壤盐分预测模型中基于盐分指数变量组的BNN模型效果最佳，R^2为0.651，RMSE为1.947；综合对比下，BNN模型的预测效果最好，可用于研究区土壤盐渍化预测。（2）银川平原主要是以非盐渍化和轻度盐渍化为主，0~20 cm土壤重度盐渍化及盐土共占总面积的11.59%，20~40 cm土壤重度盐渍化及盐土共占总面积的7.04%，20~40 cm土壤盐渍化程度较0~20 cm土壤盐渍化轻。

关 键 词：机器学习；土壤盐分预测；贝叶斯神经网络；银川平原

文章编号：1000－6060（2023）01－0103－12（0103～0114）

土壤是人类进行社会生产生活的基本条件，土壤退化牵制着农业的发展，监测土壤性质是一个重要过程，对如何可持续利用起着重要作用。全球六大洲的土壤中均有盐渍化现象发生（除南极洲未有调查数据外），并以每年约1.5×10^4 km^2的速度增加[1]。盐分预测指的是将某些指标与其相对应的地表信息间构建关系模型，众多学者对盐渍化预测展开了相关研究[2-5]。徐红涛等[6]、姜红等[7]对新疆不同区域进行土壤盐渍化预测模型的构建，研究表明机器学习的方法使得模型的预测精度高于传统的回归法预测模型。马国林等[8]基于Sentinel-2A影像，使用XGBoost选择有效变量后结合机器学习进行土壤盐分反演模型训练，模型预测效果明显提高，使得盐渍化制图精度得到提升。杨厚翔等[9]选取多个影响盐碱化的因子，借助BP网络模型对黑龙江省地区进行盐碱化危险度监测，整体准确度高于95%。刘全明等[10]利用微波雷达数据对盐分进行反演，所得BP神经网络（BPNN）模型能够在短时间内进行大范围区域的盐分监测且优于多元线性回归模型和地理加权回归模型。杨练兵等[11]在BPNN模型基础上使用遗传算法优化输入参数子集和隐含层神经元数量并对模型的初始权重进行优化，确定了适用于区域的反演模型，使得反演结果中土壤盐分含量的均质性最好。BPNN模型的使用，提升了人们对于土壤盐渍化变化及预测的精度，丰富了盐渍化研究在时空方面的意义。

人工神经网络在预测方面已经有了较多的应用成果，章龙管等[12]将贝叶斯网络应用于施工风险管理中，结果证明该方法对于施工风险及发展趋势能够有效预测，可为现场施工防范管理提供理论指导。毕春光等[13]对玉米病害进行预警研究，证明贝叶斯神经网络（BNN）模型较BP神经模型的精准度提升了5.49%。将BNN模型应用到土壤盐渍化预测中，探讨不同模型对土壤盐分的拟合效果，本文以银川平原为研究区，基于遥感指数和盐渍化影响因子，结合实测土壤盐分数据构建预测模型，选出最

收稿日期：2022-06-11； **修订日期：**2022-07-17
基金项目：国家自然科学基金项目（42061047，42067003）；宁夏回族自治区重点研发计划项目（2021BEG03002）资助
作者简介：魏慧敏（1998-），女，硕士研究生，主要从事遥感监测与分析研究. E-mail: weihm09@163.com
通讯作者：贾科利（1975-），男，博士，教授，主要从事3S与土地利用研究. E-mail: jiakl@nxu.edu.cn

优模型，预测研究区不同土壤深度的盐分分布，为银川平原土壤监测和盐渍化防治提供理论支持。

1 研究区概况

银川平原(37.83°~39.38°N，104.28°~107.65°E)西、北、东三面被腾格里沙漠、乌兰布和沙漠、毛乌素沙漠环绕，南部与黄土高原接壤。银川平原属于温带干旱区，多年平均气温5~9 ℃，年平均降水量约200 mm，蒸降比约10:1，年日照时数2500~3100 h，年太阳辐射5800×10^6~6100×10^6 J·m^{-2}。银川平原主要是由冲洪积平原构成，地势西南高东北低。黄河贯穿银川平原，优越的地理位置使得平原能够充分发展水利灌溉，灌溉渠系众多，农牧业发达[14]，主要作物有水稻、小麦、玉米。灌淤土和草甸土是主要的农业土壤，熟化程度较高。地下水矿化度为0~6 g·L^{-1}，平原地区地下水埋深较浅，土壤盐分累积较重。

2 数据与方法

2.1 土壤样本采集与盐分测定

本文大致以5 km×5 km的格网设置采样点(图1)，于2021年3月7—28日采样。采样时每个样点采用梅花形采样法，在采样点30 m×30 m范围内采集0~20 cm和20~40 cm土壤混匀，运用四分法留约500 g土壤装进密封袋带回，并用GPS记录采样点的坐标信息及周围环境信息，共采集197个样本点。采集的土壤样本剔除杂质后自然风干，研磨后过2 mm筛子，以水土比5:1配置提取液[15]，每个样品配置3组作为对照组，每个样点取3组数值的平均值作为样点最终pH和电导率，运用电导率法[16]计算土壤样本的含盐量。

$$S = (\mathrm{EC} + 41.2543) \times 5 \div 2120.76 \qquad (1)$$

式中：S为土壤含盐量(g·kg^{-1})；EC为测量所得土壤提取液电导率(S·m^{-1})。

对测量结果进行分析，删除异常样本点后获得166个样点供用。将样本的盐渍化程度分为5级[17]，样本的统计性描述如表1所示。0~20 cm土壤样本中共有91个盐渍化样本，约占总样本的54.82%。20~40 cm土壤样本共有65个盐渍化样本，约占总样本的39.16%。各层样本土壤含盐量总体变异系数均超过132%，变异性较强，表明样本离散程度较高，具有普适性。

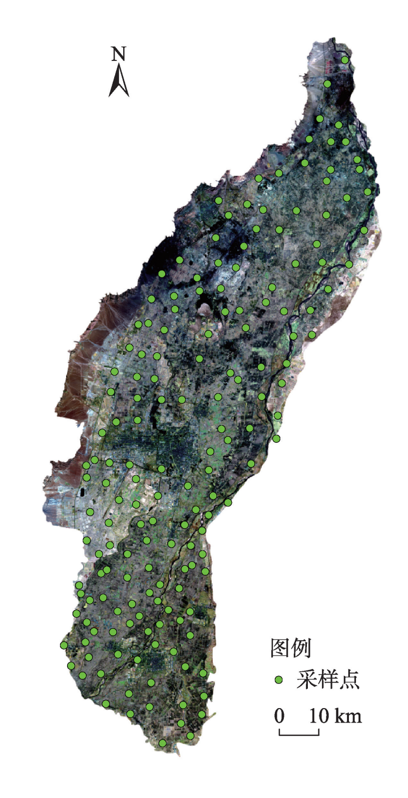


图1 银川平原采样点分布

Fig. 1 Distribution of sampling points in Yinchuan Plain

将土壤样本按含盐量由低到高排序，依据模型训练集和结果验证集2:1的比例等间隔选择110个样本进行模型训练，56个样本作为验证集用于模型验证[18]。

2.2 遥感影像获取与处理

研究选用的Landsat 8 OLI遥感影像来源于美国地质调查局官网(https://earthexplorer.usgs.gov/)，行列号为129/033和129/034，空间分辨率为30 m，成像时间为2021年3月17日，与采样时间同期。利用ENVI 5.3对所选的影像进行大气校正等预处理流程后，进行各类盐分指数的计算和提取[19]，如表2所示。

2.3 土壤盐渍化预测模型

2.3.1 盐渍化影响因子组参数选择 依据已有学者对土壤盐渍化成因条件的相关研究[26-27]，结合银川平原实际情况，选择8个环境变量和3个人为活动变量(表3)。其中，土壤pH和含水率结果由采样点数据计算后进行Kriging空间插值而得，设置输出的分辨率为30 m；地下水数据来源参考文献[17]，将数据处理为30 m分辨率的栅格数据；银川平原数字高程模型从地理空间数据云平台(http://www.gs-

表1　银川平原土壤样本统计

Tab. 1　Statistics of soil sample in Yinchuan Plain

深度/cm	程度	含盐量/g·kg^{-1}	样本数	含盐量均值/g·kg^{-1}	含盐量最大值/g·kg^{-1}	含盐量最小值/g·kg^{-1}	变异系数/%	总样本变异系数/%
0~20	非盐渍化	<1	75	0.510	0.983	0.097	41.295	172.327
	轻度盐渍化	1~2	32	1.522	1.946	1.058	18.823	
	中度盐渍化	2~4	20	2.819	3.667	2.003	20.119	
	重度盐渍化	4~6	12	4.734	5.803	4.122	11.538	
	盐土	>6	27	15.104	43.725	6.676	59.548	
20~40	非盐渍化	<1	101	0.566	0.997	0.308	30.428	132.278
	轻度盐渍化	1~2	34	1.471	1.979	1.020	17.701	
	中度盐渍化	2~4	23	2.674	3.752	2.015	19.605	
	重度盐渍化	4~6	4	4.387	5.276	4.038	13.545	
	盐土	>6	4	9.872	17.929	6.274	55.894	

表2　盐分指数计算公式

Tab. 2　Calculation formulas of salt indices

盐分指数	计算公式	参考文献
SI-T	100(Red−NIR)	刘旭辉等[19]
NDSI	(Red−NIR)/(Red+NIR)	Amal 等[20]
S1	Blue/Red	Sahana 等[21]
S2	(Blue−Red)/(Blue+Red)	Nguyen 等[22]
S3	(Green×Red)/Blue	Nguyen 等[22]
S4	$\sqrt{Blue \times Red}$	樊彦国等[23]
S5	(Blue×Red)/Green	孙亚楠等[24]
S6	(Red×NIR)/Green	孙亚楠等[24]
SI1	$\sqrt{Green \times Red}$	樊彦国等[23]
SI2	$\sqrt{Green^2 + Red^2 + NIR^2}$	樊彦国等[23]
SI3	$\sqrt{Green^2 + Red^2}$	赵巧珍等[25]

注：Blue、Green、Red、NIR分别为蓝波段(b2)、绿波段(b3)、红波段(b4)、近红外波段(b5)的光谱反射率。下同。

cloud.cn/)获取，空间分辨率为30 m；地表温度、增强植被指数和水体指数通过Landsat 8 OLI影像在ENVI 5.3软件中计算获得；土地利用数据从中国科学院资源环境科学与数据中心(http://www.resdc.cn/)获取，强度指数采用庄大方[28]利用分级方法，土地利用强度依据不同的土地类型分别确定为不同的值，未利用地、草地、林地、耕地、居民用地、水体分别赋值为1、2、2、3、4、1。

2.3.2　*数据处理*　为消除数据获取过程中人为或者自然因素导致数据不标准的问题，先将数据进行归一化处理以消除不同指标之间维度的影响。归一化公式为：

$$A_* = \frac{A - A_{min}}{A_{max} - A_{min}} \quad (1)$$

式中：A_{max} 为样本数据的最大值；A_{min} 为样本数据的最小值；A 为预警因子数据集；A_* 为归一化后的数据。

2.3.3　*预测模型构建*

(1) 支持向量机(SVM)

SVM模型基于结构风险最小原理，依靠有限的样本来检索全局最优解，拥有对未知点有较好的泛化效果的优点[29]，规避小样本学习和局部极值出现的缺点。本文选用RBF(Radial basis function)为SVM中核函数类型，惩罚参数(c)与核参量(g)由样本进行测试计算得到。

(2) BP神经网络(BPNN)

BPNN模型基于反向传播误差的方式对数据集进行训练，达到误差最小化的目的[30]。该模型拥有较强的自适应和自学习能力，可精确逼近任意的非线性关系[31]。本文基于MATLAB 2020软件，将盐分指数和影响因子作为模型的输入数据，实测盐分数据作为输出数据，创建研究区土壤盐分的BPNN预测模型。经多次训练，在参数一致的情况下，将模型的输出值和实际值进行拟合，以误差作为选择隐含层节点的标准，确定最佳训练效果的隐含层节点。模型训练目标的最小误差设置为0.00001，训练次数设置为1000次。

(3) 贝叶斯神经网络(BNN)

BNN模型通过为神经网络的权重分配引入随机性进行正则化，也相当于任意权重的多种神经网络的预测[32-33]。与传统BPNN模型不同的是贝叶斯估计得到的是节点权重的后验分布概率 $P(w|D)$，并

表3 土壤盐渍化影响因子

Tab. 3 Influence factors of soil salinization

变量类型	参数名称	描述
环境变量	土壤pH	土壤酸碱性
	土壤含水率/%	土壤绝对含水量
	高程/m	地区的海拔高度
	增强植被指数	农作物长势和营养信息
	水体指数	地区水体信息
	地表温度/℃	地区土壤温度
	地下水埋深/m	地区潜水的埋藏深度
	地下水矿化度/g·L^{-1}	区域潜水中可溶性盐类的质量
人为活动变量	土地利用强度	土地资源的利用效率
	人口密度/人·km^{-2}	每平方千米的人口平均数
	生产总值/元	生产活动的最终成果

不是固定的值,因此可以为神经网络的预测添加随机性。由给定数据输出预测值的分布,公式为:

$$P(n|m) = E_{P(w|D)}\left[P(n|m,w)\right] \tag{2}$$

式中:给定数据、预测值分别为m和n;w为每个节点的权重;D为训练集。下同。

所得到的预测值y是$P(w|D)$中的所有可能的预测值的期望。依据贝叶斯理论,样本数据决定$P(D)$。

$$P(w|D) = \frac{P(w,D)}{P(D)} = \frac{P(D|w)P(w)}{P(D)} \tag{3}$$

根据参考文献[34],获得一个高斯先验的$P(w)$:

$$P(w) = \prod_j \pi N\left(w_j|0,\sigma_1^2\right) + (1-\pi)N\left(w_j|0,\sigma_2^2\right) \tag{4}$$

式中:w_j为网络的第j个权重;σ_1^2、σ_2^2为先验的第一个和第二个混合成分的方差,且$\sigma_1 > \sigma_2$;$N\left(w_j|0,\sigma_1^2\right)$为$w_j$以均值为0评价的高斯密度。

依据所给的先验分布,运用贝叶斯定理得到后验概率,若某参数能够使后验概率达到最大化,则该参数为模型的最佳参数。BNN模型能够克服过拟合及出现局部最小化的现象。

2.3.4 模型评价指标 为量化土壤盐分预测模型的预测效果,本文选择决定系数(R^2)和均方根误差(RMSE)2个常用指标来进行综合评价。R^2的值越高,RMSE的值越低,说明模型拟合效果越好。

3 结果与分析

3.1 银川平原土壤盐渍化空间分布特征分析

使用ArcGIS的空间分析工具对166个采样点的土壤含盐量进行插值,根据盐渍化程度划分为5个等级,得到银川平原不同深度土壤含盐量的空间分布特征(图2)。由图2可见,0~20 cm重度盐渍化土壤及盐土主要分布在银川平原北部地区,南部地区盐渍化程度较轻。20~40 cm主要以非盐渍化和轻度盐渍化土壤为主,银川平原西北部土壤盐渍化现象严重,部分区域土壤含盐量大于6 g·kg^{-1}。整体上研究区0~20 cm和20~40 cm土壤盐渍化呈现出南低北高的趋势,重度盐渍化土壤和盐土主要分布在石嘴山地区。

3.2 不同预测模型对比分析

3.2.1 相关性分析 对盐渍化影响因子、盐分指数与实测土壤盐分进行相关性分析,所得结果如表4所示。由表4可见,影响因子中,0~20 cm土壤除土壤pH、土壤含水率、人口密度外,其余因子与含盐量具有较强的显著性;20~40 cm土壤含盐量与土地利用强度、增强植被指数、水体指数、地表温度、地下水埋深、地下水矿化度在0.01水平上具有极强的显著性。盐分指数中,0~20 cm土壤含盐量与S1、S2、S3、S6、SI2、NDSI、SI-T的相关性均通过0.05检验。20~40 cm土壤含盐量与S1、S2、S4、S5、SI3、NDSI、SI-T的相关性均通过0.05检验。故0~20 cm土壤选择通过0.01显著性水平检验的影响因子和盐分指数分别作为输入层参数进行建模,20~40 cm土壤选择通过0.05显著性水平检验的影响因子和盐分指数分别作为输入层参数进行建模。

3.2.2 模型构建与对比分析 选择SVM、BPNN和BNN模型,以土壤含盐量为输出层,相关性强的影

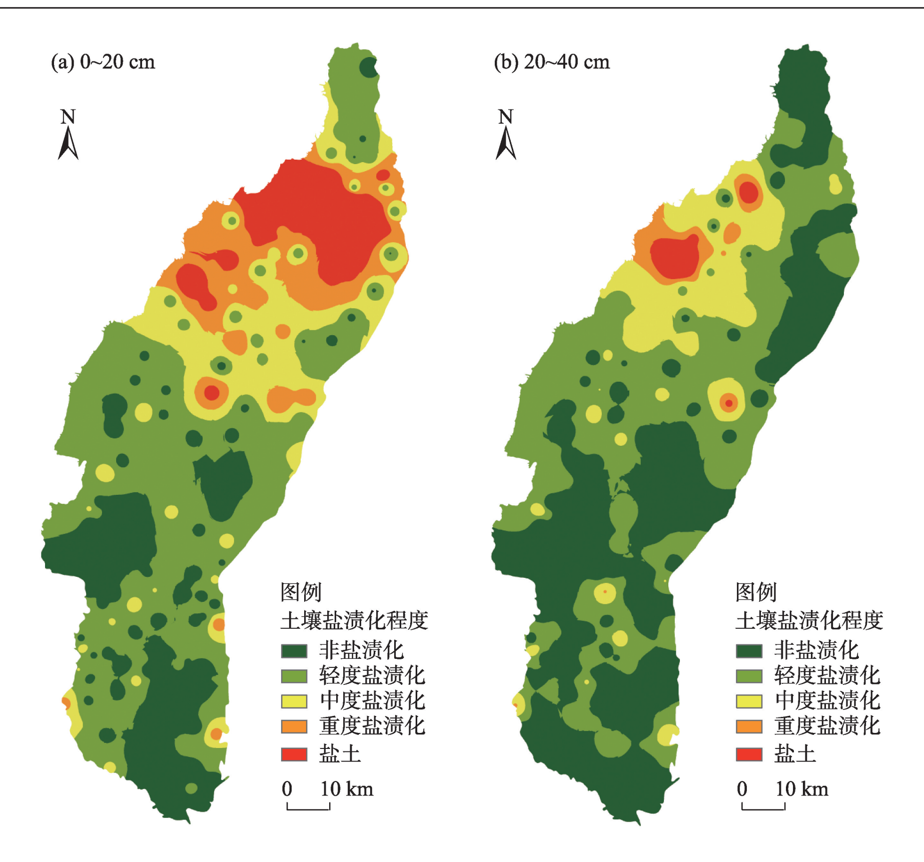


图2 土壤盐分空间分布

Fig. 2 Spatial distribution of soil salinity

响因子和盐分指数为输入层数据，对0~20 cm和20~40 cm土壤盐分含量进行模拟，结果如表5所示。以验证集样本对模型进行验证，结果如图3。由表5可见，0~20 cm土壤影响因子变量组的3个模型R^2均大于0.76，其中BNN模型建模效果最好，BPNN模型次之，SVM模型最差。由验证结果表明，BNN模型的验证效果与SVM模型和BPNN模型相比结果最佳，其模拟值与实测值的R^2值最大(图3)。盐分指数组的R^2整体较低，但均通过0.01显著性检验。其中，该分组中BNN模型的验证集RSME为最低，BPNN模型拟合误差最大。对比6个模型的验证效果，发现基于影响因子组建立的BNN模型是0~20 cm土壤盐分预测最佳模型，故选择此模型对银川平原0~20 cm土壤盐分进行预测。

20~40 cm土壤盐分预测模型中(表5)，影响因子组模型验证集的RMSE均大于2，BNN模型的R^2大于SVM和BPNN模型，BNN模型的RMSE小于SVM和BPNN模型，因此BNN模型检验结果效果优于SVM和BPNN模型(图3)。盐分指数组验证集的R^2相比于影响因子组有明显提升，盐分指数组BNN模型是3个模型中误差最小、拟合程度最佳的模型，可以将该模型用于研究区20~40 cm土壤盐分预测。

3.3 银川平原土壤盐分预测

利用经过训练的BNN模型对整个银川平原0~20 cm、20~40 cm土壤进行盐分预测，而后按照盐渍化程度对结果划分等级，得到银川平原土壤盐分预测等级图(图4)，由图4可知，0~20 cm土壤中盐土主要分布在银川平原北部的大武口区和平罗县，重度盐渍化和中度盐渍化土壤主要分布在盐土四周及平原东部地区，非盐渍化和轻度盐渍化土壤主要分布在平原南部地区。0~20 cm土壤主要是以非盐渍化土壤为主，盐土分布于大武口区及平罗县西部地区，重度盐渍化土壤零星分布在银川平原内。根据像元计算得到，0~20 cm土壤主要是以非盐渍化土壤为主，占银川平原总面积的33.29%，其次为轻度盐渍化土壤，面积占比为31.86%，中度盐渍化和重度盐渍化土壤面积占比分别为23.26%和8.07%，盐土面积占比(3.52%)不足5%。20~40 cm土壤中非盐渍化土壤和轻度盐渍化土壤面积占比分别为50.54%和30.07%，中度盐渍化土壤面积占比为

表4 模型参数与土壤盐分的相关性

Tab. 4 Correlation between model parameters and soil salinity

参数类型	参数	与0~20 cm土壤相关系数(r_1)	与20~40 cm土壤相关系数(r_2)
盐渍化影响因子	土壤pH	0.034	0.030
	土壤含水率	0.103	0.102
	高程	−0.253**	−0.114
	土地利用强度	−0.403**	−0.255**
	增强植被指数	−0.580**	−0.311**
	水体指数	0.511**	0.271**
	地表温度	−0.445**	−0.119**
	地下水埋深	−0.477**	−0.204**
	地下水矿化度	0.768**	0.254**
	人口密度	−0.023	0.048
	生产总值	−0.173*	−0.025
盐分指数	S1	0.492**	0.402**
	S2	0.480**	0.388**
	S3	−0.230**	−0.054
	S4	0.049	0.210**
	S5	0.067	0.229**
	S6	−0.309**	−0.093
	SI1	−0.032	0.136
	SI2	−0.130*	0.069
	SI3	−0.043	0.127*
	NDSI	0.325**	0.204**
	SI-T	0.347**	0.182*

注:*、**分别表示$P<0.05$、$P<0.01$水平上的显著性检验。

表5 基于不同变量组的机器学习模型

Tab. 5 Machine learning models based on different variable groups

变量组	模型类别	建模集(n=110)		验证集(n=56)	
		R^2	RMSE	R^2	RSME
影响因子组(0~20 cm)	SVM	0.766	3.397	0.542	3.266
	BPNN	0.773	2.638	0.608	3.081
	BNN	0.797	2.751	0.618	2.986
盐分指数组(0~20 cm)	SVM	0.319	5.945	0.213	4.155
	BPNN	0.311	5.878	0.206	4.405
	BNN	0.336	5.715	0.250	3.890
影响因子组(20~40 cm)	SVM	0.368	1.327	0.221	2.401
	BPNN	0.314	1.353	0.331	2.221
	BNN	0.425	1.076	0.486	2.026
盐分指数组(20~40 cm)	SVM	0.440	1.032	0.436	2.079
	BPNN	0.348	1.086	0.545	2.033
	BNN	0.442	1.006	0.651	1.947

注:SVM、BPNN、BNN分别为支持向量机、BP神经网络和贝叶斯神经网络;R^2为决定系数;RMSE为均方根误差。

12.35%,重度盐渍化土壤和盐土面积占比(分别为4.89%、2.15%)均小于5%。

为进一步验证预测结果与实验室结果的一致性,将预测结果与实测盐分插值结果进行栅格对比,统计相同等级像元数据(表6)。由表6可知,0~20 cm和20~40 cm土壤盐渍化预测所得的盐渍化程

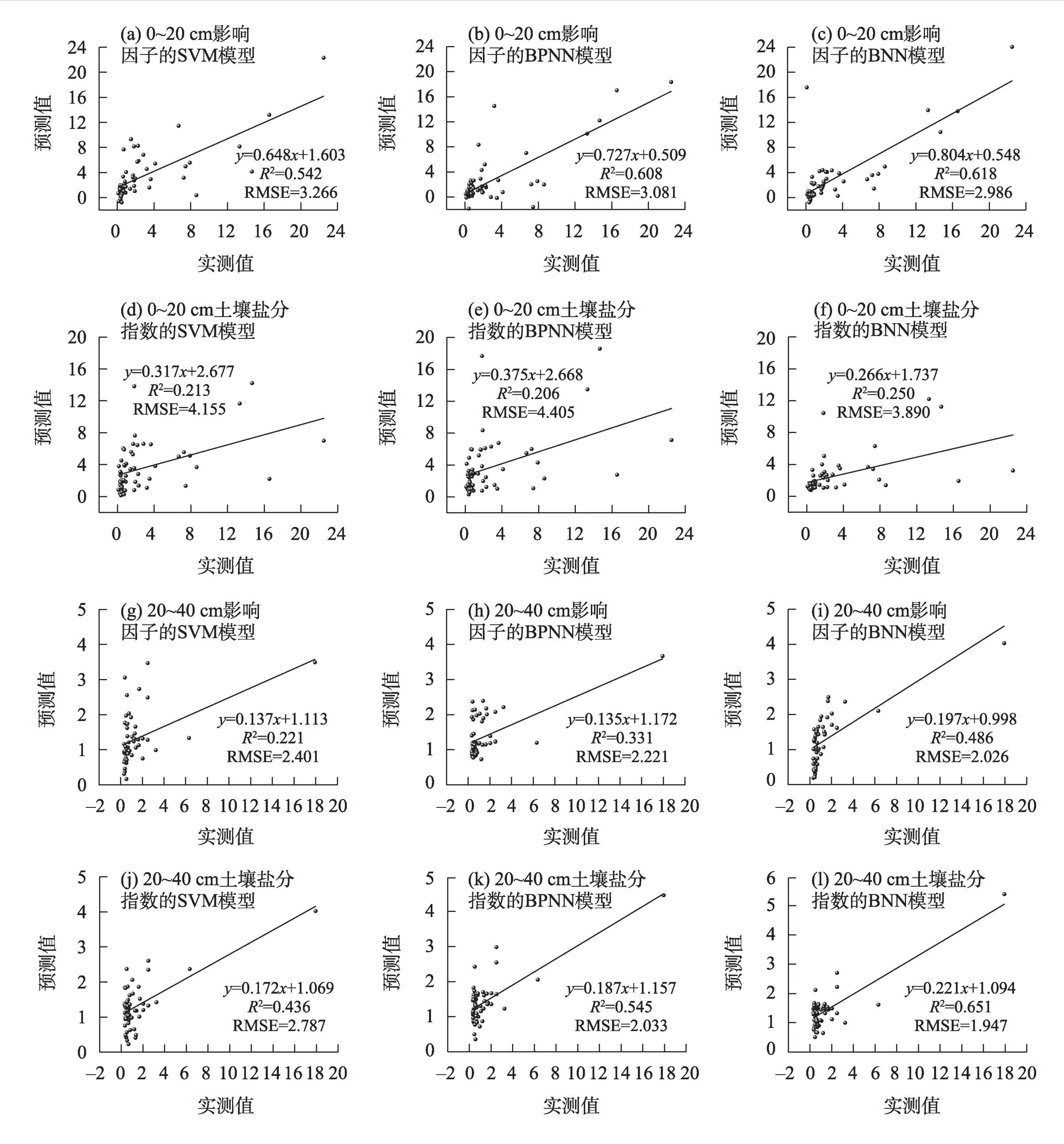


图3 土壤盐分与模型预测值的散点图

Fig. 3 Scatter plot of soil salinity and model predictions value

度与实测含盐量插值相同等级像元个数的相同率均大于73%，表明预测模型的结果与实测数据间具有较好的一致性，所选模型可达到检验要求并用于研究区的土壤盐渍化预测。

4 讨 论

使用3种机器学习的方法建立了对于不同深度的土壤盐分预测模型，发现不同深度的土壤对同种变量在相同模型下的建模效果差异较大，对所有结果进行精度评价后，选择适用不同深度土壤的最佳模型进行土壤盐分预测。

本文对多个影响因子与实测值进行Pearson相关性分析，发现敏感因子为地下水矿化度、地下水埋深、土地利用强度等多个因子，与杨思存等[35]、苏春利等[36]所得结论相符，后续可用于挖掘因素与土壤盐分的隐含信息。0~20 cm土壤盐分与各影响因子的相关性明显高于20~40 cm土壤，这与刘继龙等[37]研究结果一致，可能是土壤盐分受到外界影响程度随深度不同而改变，人们所进行的农业活动主

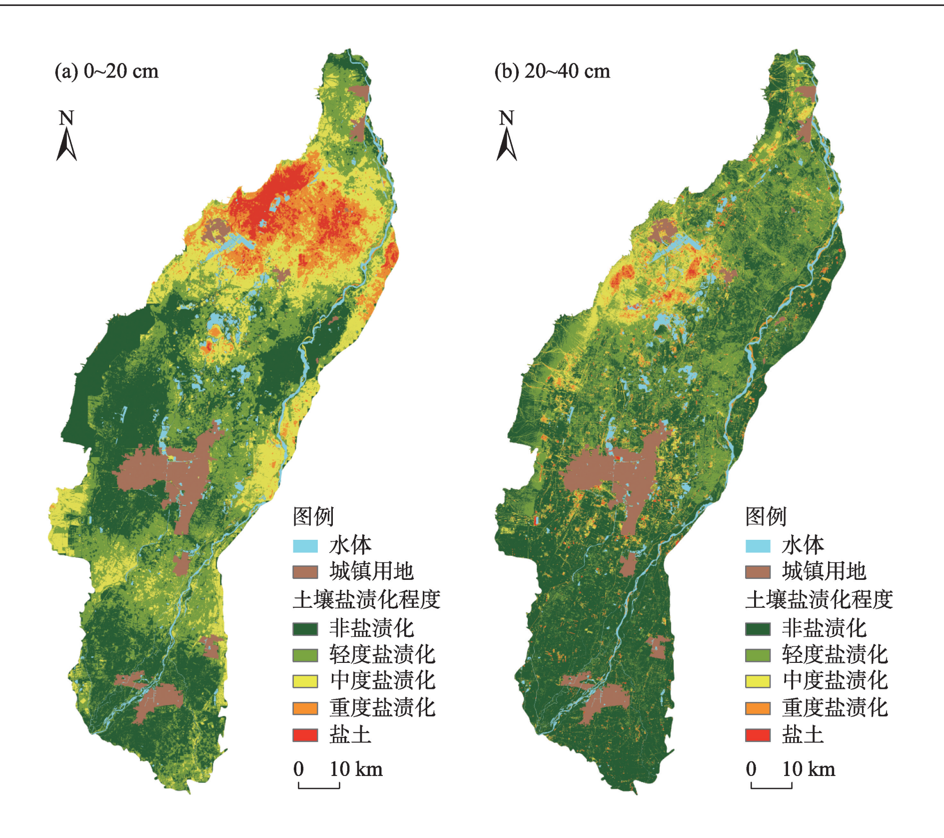


图4 土壤含盐量预测等级分布

Fig. 4 Grade distribution of soil salinity inversion

表6 实测插值与预测像元统计

Tab. 6 Measured interpolation and predictive pixel statistics

深度/cm	盐渍化程度	实测插值像元数	预测和实测插值相同等级个数	相同率/%
0~20	非盐渍化	3922460	3221641	82.13
	轻度盐渍化	1925470	1409137	73.18
	中度盐渍化	779011	648843	83.29
	重度盐渍化	764161	602002	78.78
	盐土	507095	434657	85.72
20~40	非盐渍化	4279527	3613169	84.43
	轻度盐渍化	2413166	2046062	84.79
	中度盐渍化	722601	538546	74.53
	重度盐渍化	311423	274697	88.21
	盐土	171493	144933	84.51

要是在0~20 cm土壤，对20~40 cm土壤的干预性较小。

边慧芹等[38]、何宝忠等[39]对于不同地区的盐渍化进行研究，表明盐分指数在监测应用方面具有一定的优势，选择多种盐分指数与实测盐分进行相关性分析，选取相关性较高的盐分指数作为模型参数进行预测模型训练。本文采样时正值春季，气候干燥，蒸发强烈，土壤盐分随水分的蒸发而在地表集聚，使0~20 cm土壤含盐量高于20~40 cm土壤，造成盐分指数与两层土壤盐分的相关性差异。土壤是一个复杂的综合体，常规的线性回归模型难以达到较好的精度，机器学习能够较好地克服这个难题，许多研究者将机器学习用于预测盐分，所得模型精度高于线性模型[5,40-41]，BNN算法可运用较少的数据得到稳定的模型和各层参数的分布，达到较好的泛化目的，可为盐分预测提供新思路。本文基于2个

变量组，对3种不同机器学习方法进行建模比较，发现BNN模型精度优于BPNN模型和SVM模型。且0~20 cm土壤盐分预测整体效果优于20~40 cm深度的模型效果，与杨宁等[42]所得结论基本一致。

本文通过对比不同深度和不同输入参数下多个模型的预测效果，选择最佳模型进行预测。同一地区不同季节不同年份的环境差异较大，春季银川平原土壤中的盐分不断向地表层汇聚，银川平原主要为耕地，农民在耕种前会对农田进行浸灌以此来降低农田土壤盐渍化程度，后续研究应该考虑不同季节的环境变化，选择多年多季节数据进行预测模型的训练及验证。两组输入参数对土壤盐渍化程度的预测精度不同，在下一步工作中将对影响因子与盐分指数进行组合参数，进行区域土壤盐分的预测模型训练选优，提高模型的预测精度。另一方面，不同地区土壤盐渍化驱动机制不同，本文所建立的预测模型能否运用到其他地区还需进行下一步的研究和验证。

5　结　论

以银川平原为研究区，将盐渍化影响因子和盐分指数作为输入参数构建盐分预测模型，对土壤0~20 cm及20~40 cm土壤进行预测，得到以下结论：

（1）通过土壤盐分预测模型对比发现，BNN模型进行土壤盐分预测效果最优，影响因子对0~20 cm土壤盐分预测的影响较大；盐分指数对20~40 cm土壤盐分预测的影响较大。对比分析所选算法的不同变量建模和验证效果，表明BNN模型是效果最好的预测模型，神经网络的引入对模型的训练有一定的优势。

（2）根据预测结果可知，银川平原0~20 cm土壤主要是以非盐渍化和轻度盐渍化为主，面积占比分别为33.29%、31.86%，重度盐渍化及盐土面积占比为11.59%。20~40 cm土壤中非盐渍化和轻度盐渍化土壤面积占比分别为50.54%和30.07%，重度盐渍化及盐土面积占比为7.04%。20~40 cm土壤盐渍化程度较0~20 cm土壤盐渍化轻。

参考文献(References)

[1] 杨真, 王宝山. 中国盐渍土资源现状及改良利用对策[J]. 山东农业科学, 2015, 47(4): 125−130. [Yang Zhen, Wang Baoshan. Present status of saline soil resources and countermeasures for improvement and utilization in China[J]. Shandong Agricultural Sciences, 2015, 47(4): 125−130.]

[2] 翁永玲, 戚浩平, 方洪宾, 等. 基于PLSR方法的青海茶卡−共和盆地土壤盐分高光谱遥感反演[J]. 土壤学报, 2010, 47(6): 1255−1263. [Weng Yongling, Qi Haoping, Fang Hongbin, et al. PLSR-based hyperspectral remote sensing retrieval of soil salinity of Chaka-Gonghe Basin in Qinghai Province[J]. Acta Pedologica Sinica, 2010, 47(6): 1255−1263.]

[3] 王爽, 丁建丽, 王璐, 等. 基于地表光谱建模的区域土壤盐渍化遥感监测研究[J]. 干旱区地理, 2016, 39(1): 190−198. [Wang Shuang, Ding Jianli, Wang Lu, et al. Remote sensing monitoring of soil salinization based on surface spectral modeling[J]. Arid Land Geography, 2016, 39(1): 190−198.]

[4] El-Horiny M M. Mapping and monitoring of soil salinization using remote sensing and regression techniques: A case study in the Bahariya depression, western desert, Egypt[C]//2019 IEEE International Geoscience and Remote Sensing Symposium. Yokohama: Institute of Electrical and Electronic Engineers, 2019.

[5] 陈红艳, 赵庚星, 陈敬春, 等. 基于改进植被指数的黄河口区盐渍土盐分遥感反演[J]. 农业工程学报, 2015, 31(5): 107−114. [Chen Hongyan, Zhao Gengxing, Chen Jingchun, et al. Remote sensing inversion of saline soil salinity based on modified vegetation index in estuary area of Yellow River[J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(5): 107−114.]

[6] 徐红涛, 陈春波, 郑宏伟, 等. 集成建模变量优选和参数学习的SVR盐渍化监测[J]. 遥感技术与应用, 2021, 36(1): 176−186. [Xu Hongtao, Chen Chunbo, Zheng Hongwei, et al. SVR salinization monitoring based on integrated feature subset selection and model parameter learning[J]. Remote Sensing Technology and Application, 2021, 36(1): 176−186.]

[7] 姜红, 玉素甫江·如素力, 热伊莱·卡得尔, 等. 基于神经网络模型的干旱区绿洲土壤盐渍化评价分析[J]. 地球信息科学学报, 2017, 19(7): 983−993. [Jiang Hong, Rusuli Yusufujiang, Kadeer Reyilai, et al. Evaluation and analysis of soil salinization in the arid zones based on neural network model[J]. Journal of Geo-information Science, 2017, 19(7): 983−993.]

[8] 马国林, 丁建丽, 韩礼敬, 等. 基于变量优选与机器学习的干旱区湿地土壤盐渍化数字制图[J]. 农业工程学报, 2020, 36(19): 124−131. [Ma Guolin, Ding Jianli, Han Lijing, et al. Digital mapping of soil salinization in arid area wetland based on variable optimized selection and machine learning[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(19): 124−131.]

[9] 杨厚翔, 雷国平, 徐秋, 等. 基于危险度与风险格局的土地盐碱化监测区优先级评价[J]. 农业工程学报, 2019, 35(7): 238−246. [Yang Houxiang, Lei Guoping, Xu Qiu, et al. Priority evaluation of land salinization monitoring area based on danger degree and risk pattern[J]. Transactions of the Chinese Society of Agricultural En-

gineering, 2019, 35(7): 238-246.]

[10] 刘全明, 成秋明, 王学, 等. 河套灌区土壤盐渍化微波雷达反演[J]. 农业工程学报, 2016, 32(16): 109-114. [Liu Quanming, Cheng Qiuming, Wang Xue, et al. Soil salinity inversion in Hetao irrigation district using microwave radar[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(16): 109-114.]

[11] 杨练兵, 郑宏伟, 罗格平, 等. 基于遗传算法优化BP神经网络的土壤盐渍化反演[J]. 地理与地理信息科学, 2021, 37(2): 12-21, 37. [Yang Lianbing, Zheng Hongwei, Luo Geping, et al. Retrieval of soil salinity content based on BP neural network optimized by genetic algorithm[J]. Geography and Geo-information Science, 2021, 37(2): 12-21, 37.]

[12] 章龙管, 刘绥美, 李开富, 等. 基于故障树与贝叶斯网络的地铁盾构施工风险预测[J]. 现代隧道技术, 2021, 58(5): 21-29, 55. [Zhang Longguan, Liu Suimei, Li Kaifu, et al. Prediction of shield construction risks in subway tunneling based on fault tree and Bayesian network[J]. Modern Tunneling Technology, 2021, 58(5): 21-29, 55.]

[13] 毕春光, 王金龙, 胡楠, 等. 基于贝叶斯神经网络的玉米病害预警模型[J]. 吉林农业大学学报, 2021, 43(2): 189-195. [Bi Chunguang, Wang Jinlong, Hu Nan, et al. A Bayesian neural network-based early warning model for maize diseases[J]. Journal of Jilin Agricultural University, 2021, 43(2): 189-195.]

[14] 靳晓辉, 樊玉苗, 段浩, 等. 银川平原地下水位对黄河流域水量统一调度的时空响应分析[J]. 水资源与水工程学报, 2021, 32(4): 45-51. [Jin Xiaohui, Fan Yumiao, Duan Hao, et al. Temporal and spatial response of groundwater depth in Yinchuan Plain to the integrated water regulation of the Yellow River[J]. Journal of Water Resources and Water Engineering, 2021, 32(4): 45-51.]

[15] 曹肖奕, 丁建丽, 葛翔宇, 等. 基于不同卫星光谱模拟的土壤电导率估算研究[J]. 干旱区地理, 2020, 43(1): 172-181. [Cao Xiaoyi, Ding Jianli, Ge Xiangyu, et al. Estimation of soil conductivity based on spectral simulation of different satellites[J]. Arid Land Geography, 2020, 43(1): 172-181.]

[16] 贾萍萍, 尚天浩, 张俊华, 等. 利用多源光谱信息反演宁夏银北地区干湿季土壤含盐量[J]. 农业工程学报, 2020, 36(17): 125-134. [Jia Pingping, Shang Tianhao, Zhang Junhua, et al. Inversion of soil salinity in dry and wet seasons based on multi-source spectral data in Yinbei area of Ningxia, China[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(17): 125-134.]

[17] 武丹, 贾科利, 张晓东, 等. 基于异质SVM神经网络的土壤盐渍化灾害预测模型[J]. 水文地质工程地质, 2018, 45(5): 143-149, 157. [Wu Dan, Jia Keli, Zhang Xiaodong, et al. Soil salinization disaster prediction model based on heterogeneous SVM neural network[J]. Hydrogeology & Engineering Geology, 2018, 45(5): 143-149, 157.]

[18] 张智韬, 魏广飞, 姚志华, 等. 基于无人机多光谱遥感的土壤含盐量反演模型研究[J]. 农业机械学报, 2019, 50(12): 151-160. [Zhang Zhitao, Wei Guangfei, Yao Zhihua, et al. Soil salt inversion model based on UAV multispectral remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(12): 151-160.]

[19] 刘旭辉, 白云岗, 柴仲平, 等. 基于多光谱遥感的典型绿洲棉田春季土壤盐分反演及验证[J]. 干旱区地理, 2022, 45(4): 1165-1175. [Liu Xuhui, Bai Yungang, Cai Zhongping, et al. Inversion and validation of soil salinity based on multispectral remote sensing in typical oasis cotton field in spring[J]. Arid Land Geography, 2022, 45(4): 1165-1175.]

[20] Amal A, Lalit K, Youself Y A. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region [J]. Geoderma, 2014, 230-231(7): 1-8.

[21] Sahana M, Rehman S, Patel P P, et al. Assessing the degree of soil salinity in the Indian Sundarban Biosphere Reserve using measured soil electrical conductivity and remote sensing data-derived salinity indices[J]. Arabian Journal of Geosciences, 2020, 13(24): 1289, doi: s12517-020-06310-w.

[22] Nguyen K A, Liou Y A, Tran H P, et al. Soil salinity assessment by using near-infrared channel and vegetation soil salinity index derived from Landsat 8 OLI data: A case study in the Tra Vinh Province, Mekong Delta, Vietnam[J]. Progress in Earth and Planetary Science, 2020, 7(1): 1-16.

[23] 樊彦国, 张维康, 刘敬一. 基于植被指数-盐分指数特征空间的黄河三角洲盐渍化遥感监测研究[J]. 山东农业科学, 2016, 48(5): 137-141. [Fan Yanguo, Zhang Weikang, Liu Jingyi. Remote sensing monitoring model of soil salinization in the Yellow River Delta Zone based on vegetation index-salt index feature space[J]. Shandong Agricultural Sciences, 2016, 48(5): 137-141.]

[24] 孙亚楠, 李仙岳, 史海滨, 等. 基于多源数据融合的盐分遥感反演与季节差异性研究[J]. 农业机械学报, 2020, 51(6): 169-180. [Sun Ya'nan, Li Xianyue, Shi Haibin, et al. Remote sensing inversion of soil salinity and seasonal difference analysis based on multi-source data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 169-180.]

[25] 赵巧珍, 丁建丽, 韩礼敬, 等. MODIS和Landsat时空融合影像在土壤盐渍化监测中的适用性研究——以渭干河-库车河三角洲绿洲为例[J]. 干旱区地理, 2022, 45(4): 1155-1164. [Zhao Qiaozhen, Ding Jianli, Han Lijing, et al. Exploring the application of MODIS and Landsat spatiotemporal fusion images in soil salinization: A case of Weigan River-Kuqa River Delta Oasis[J]. Arid Land Geography, 2022, 45(4): 1155-1164.]

[26] 赖宁, 李新国, 阿斯耶姆·图尔迪, 等. 开都河流域下游绿洲土壤盐渍化时空变化及其成因分析[J]. 干旱区资源与环境, 2013, 27(12): 66-73. [Lai Ning, Li Xinguo, Tuerdi Asiyemu, et al. Analysis of tempo-spatial dynamic change and the cause of soil salinization in the oasis of the lower reaches of Kaidu River Basin[J]. Journal of Arid Land Resources and Environment, 2013, 27(12): 66-

73.]

[27] 庄庆威, 吴世新, 杨怡, 等. 近10年新疆不同程度盐渍化耕地的时空变化特征[J]. 中国科学院大学学报, 2021, 38(3): 341-349. [Zhuang Qingwei, Wu Shixin, Yang Yi, et al. Spatiotemporal characteristics of different degrees of salinized cultivated land in Xinjiang in the recent ten years[J]. Journal of University of Chinese Academy of Sciences, 2021, 38(3): 341-349.]

[28] 庄大方, 刘纪远. 中国土地利用程度的区域分异模型研究[J]. 自然资源学报, 1997, 12(2): 10-16. [Zhuang Dafang, Liu Jiyuan. Study on the model of regional differentiation of land use degree in China[J]. Journal of Natural Resources, 1997, 12(2): 10-16.]

[29] 汪海燕, 黎建辉, 杨风雷. 支持向量机理论及算法研究综述[J]. 计算机应用研究, 2014, 31(5): 1281-1286. [Wang Haiyan, Li Jianhui, Yang Fenglei. Overview of support vector machine analysis and algorithm[J]. Application Research of Computers, 2014, 31(5): 1281-1286.]

[30] 杨淑娥, 黄礼. 基于BP神经网络的上市公司财务预警模型[J]. 系统工程理论与实践, 2005, 25(1): 12-18, 26. [Yang Shu'e, Huang Li. Financial crisis warning model based on BP neural network [J]. Systems Engineering-Theory & Practice, 2005, 25(1): 12-18, 26.]

[31] 赵亚琴. 基于模糊神经网络的火灾识别算法[J]. 计算机仿真, 2015, 32(2): 369-373. [Zhao Yaqin. Forest fire recognition algorithm based on fuzzy neural network[J]. Computer Simulation, 2015, 32(2): 369-373.]

[32] 宿嘉颖. 贝叶斯深度网络的不确定性分析[D]. 哈尔滨: 哈尔滨工业大学, 2020. [Su Jiaying. Uncertainty analysis of Bayesian deep network[D]. Harbin: Harbin Institute of Technology, 2020.]

[33] 孙发友, 蒙祖强. 一种设备状态监测的贝叶斯正则化BP神经网络[J]. 网络安全技术与应用, 2019(10): 48-53. [Sun Fayou, Meng Zuqiang. A Bayesian regularized BP neural network for equipment condition monitoring[J]. Network Security Technology & Application, 2019(10): 48-53.]

[34] 王子涵, 杨秀芝, 段现银, 等. 基于贝叶斯神经网络的机床热误差建模[J]. 制造技术与机床, 2022(1): 141-145. [Wang Zihan, Yang Xiuzhi, Duan Xianyin, et al. Research on the thermal error modeling of machine tool based on Bayesian neural network[J]. Manufacturing Technology & Machine Tool, 2022(1): 141-145.]

[35] 杨思存, 车宗贤, 王成宝, 等. 甘肃沿黄灌区土壤盐渍化特征及其成因[J]. 干旱区研究, 2014, 31(1): 57-64. [Yang Sicun, Che Zongxian, Wang Chengbao, et al. Soil salinization and its causes in the irrigated areas along the Yellow River in Gansu Province[J]. Arid Zone Research, 2014, 31(1): 57-64.]

[36] 苏春利, 纪倩楠, 陶彦臻, 等. 河套灌区西部土壤盐渍化分异特征及其主控因素[J]. 干旱区研究, 2022, 39(3): 916-923. [Su Chunli, Ji Qiannan, Tao Yanzhen, et al. Differentiation characteristics and factors of soil salinization in the west of Hetao Irrigation Area[J]. Arid Zone Research, 2022, 39(3): 916-923.]

[37] 刘继龙, 刘璐, 马孝义, 等. 不同尺度不同土层土壤盐分的空间变异性研究[J]. 应用基础与工程科学学报, 2018, 26(2): 305-312. [Liu Jilong, Liu Lu, Ma Xiaoyi, et al. Spatial variability of soil salt in different soil layers at different scales[J]. Journal of Basic Science and Engineering, 2018, 26(2): 305-312.]

[38] 边慧芹, 王雪梅. 基于多光谱影像的干旱区绿洲耕层土壤盐分估算[J]. 干旱区资源与环境, 2022, 36(5): 110-118. [Bian Huiqin, Wang Xuemei. Estimation of soil salinity in cultivated layers of oasis in arid areas based on multispectral images[J]. Journal of Arid Land Resources and Environment, 2022, 36(5): 110-118.]

[39] 何宝忠, 丁建丽, 刘博华, 等. 渭库绿洲土壤盐渍化时空变化特征[J]. 林业科学, 2019, 55(9): 185-196. [He Baozhong, Ding Jianli, Liu Bohua, et al. Spatiotemporal variation of soil salinization in Weigan-Kuqa River Delta Oasis[J]. Scientia Silvae Sinicae, 2019, 55(9): 185-196.]

[40] 陈俊英, 姚志华, 张智韬, 等. 大田葵花土壤含盐量无人机遥感反演研究[J]. 农业机械学报, 2020, 51(7): 178-191. [Chen Junying, Yao Zhihua, Zhang Zhitao, et al. UAV remote sensing inversion of soil salinity in field of sunflower[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 178-191.]

[41] 王明宽, 莫宏伟, 陈红艳. 基于多光谱影像反演土壤盐分的建模方法研究[J]. 土壤通报, 2016, 47(5): 1036-1041. [Wang Mingkuan, Mo Hongwei, Chen Hongyan. Study on model method of inversion of soil salt based on multispectral image[J]. Chinese Journal of Soil Science, 2016, 47(5): 1036-1041.]

[42] 杨宁, 崔文轩, 张智韬, 等. 无人机多光谱遥感反演不同深度土壤盐分[J]. 农业工程学报, 2020, 36(22): 13-21. [Yang Ning, Cui Wenxuan, Zhang Zhitao, et al. Soil salinity at different depths using improved spectral index with UAV multispectral remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(22): 13-21.]

Prediction of soil salinity based on machine learning and multispectral remote sensing in Yinchuan Plain

WEI Huimin[1], JIA Keli[1], ZHANG Xu[1], ZHANG Junhua[2]

(1. College of Geographical Sciences and Planning, Ningxia University, Yinchuan 750021, Ningxia, China; 2. Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwestern China, School of Ecology and Environment, Ningxia University, Yinchuan 750021, Ningxia, China)

Abstract: Soil salinization can hinder agricultural development. In this study, the degree of regional soil salinization was obtained to provide a theoretical reference for improving agricultural land quality. Using Yinchuan Plain of China as the study area with a grid size of 5 km×5 km, the soil salinity data of 166 sampling points at different depths were obtained. Combined with the Landsat 8 OLI image corresponding to the sampling time, the salt influence factor and salt index were used as input parameters, respectively, and soil salinity at field sampling points was used as output layer parameters. Support vector machine, back propagation neural network, and Bayesian neural network (BNN) were established as soil salinity inversion models. The determination coefficient and root mean square error of the different models were compared to screen the best model. Finally, soil salinization inversion at different depths was performed in the study area. The following results were obtained: (1) In the 0−20 cm soil salinity inversion model, the BNN model based on the influence factor variable group of salinization was the best, with a coefficient of determination (R^2) and root mean square error (RMSE) of 0.618 and 2.986, respectively; the best inversion result of 20−40 cm soil salinity was the BNN model based on the salt index variable group (R^2=0.651; RMSE=1.947); the comparative analysis of the modeling and verification effects of different variables of the selected algorithms revealed that the BNN model was the best inversion model with a better fitting degree than the other two models, and the introduction of a neural network had certain advantages in the model construction. (2) Non-salinized and mildly salinized soils were the main soil types in Yinchuan Plain. Soil salinization showed a low trend in the south and a high trend in the north. The 20−40 cm soil salinization was found to be lighter than the 0−20 cm soil salinization.

Key words: machine learning; soil salinity prediction; Bayesian neural network; Yinchuan Plain